Multimodal Studies of Instructional Discourse

Brian MacWhinney
Carnegie Mellon University

Catherine Snow
Harvard University

This workshop was held at Carnegie Mellon University in Pittsburgh on December 4-5, 1999. It was organized by Brian MacWhinney (CMU) and Catherine Snow (Harvard) and was attended by 30 invited participants. The goal of the workshop was to examine the use of new multimedia technologies for creating an integrated database of studies of instructional discourse.   The specific proposal that was discussed was whether it would be possible and desirable to establish a national data exchange system for multimodal records of instructional interactions within the framework of the NSF TalkBank Project, directed by Brian MacWhinney. If such a database could be established, it would serve as a focal point for new empirical studies of education, learning, and the transition of the child to the workforce.

This proposal involved three major components: (1) an examination of the current state of technology in the multimodal study of interactions, (2) an overview of current issues in classroom and tutorial discourse, and (3) a confrontation of issues involved in data-sharing and the protection of confidentiality.   These three issues are reviewed separately in the following sections.

The Technology.

Multimedia technologies are developing quite rapidly and are becoming much easier to use. The crucial developments are:

1. **Digitization**. It is now possible for researchers to use a reasonably priced system based on digital camcorders, mini-DV, FireWire, and digitization software such as Final Cut or iMovie to produce high quality digital video that can be stored on media such as DVD-RAM.

2. **Transcription**. There are now a variety of systems that allow researchers to gain immediate access to segments of video for analysis and transcription. The TalkBank Project, which began in September, 1999, is committed to developing tools of this type. The first such tool that it has developed is a video editor based on the earlier tools of the CHILDES project. In addition, TalkBank is currently extending the Transcriber tool developed in Paris by Claude Barras. The extended form of Transcriber will allow for video transcription across all current audio and video formats.

3. **Codon**. TalkBank plans the construction of an abstract level of annotation representation called Codon. This system is being built by Steven Bird at the University of Pennsylvania. All current transcription systems can be translated into Codon. Because it can serve as an interlingua, it is then only necessary to build one translator for each coding system, rather than all the pair-wise translations that would otherwise be needed (N*[N-1]). Making the translation process automatic will make it possible to build software that is general across coding systems. In addition, Codon is XML compatible and can integrate well with tools for web-based retrieval from databases.

4. **Retrieval**. Documentation of teaching and learning often requires the linking of a wide variety of materials, including sketches, pictures, video, diaries, notes, and other documents. All of these materials can be digitized, along with supporting tutorial materials. Browsing through such folios can be done either through HTML or special purpose software.

5. **Collaborative commentary**. In early 1999, the journal *Discourse Processes* published a special issue with CD-ROM edited by Tim Koschmann. The CD-ROM contained a digitized video of a 6-minute segment from problem-based learning in a medical school context. Six researchers provided alternative interpretations of this segment. This special issue showed how multimedia can focus the research agenda through facilitating collaborative commentary. In his workshop presentation, Koschmann showed how commentary of this sort can advance our understanding of problem-based learning. Carl Frederiksen's presentation in this workshop outlined his system of discourse analysis particularly as it applied to the segment in this special issue.

The Issues

The presentations in the workshop highlighted the importance of distinguishing two alternative uses of multimodal data. The first focuses on teacher training. Multimodal data can help teachers gain a clearer understanding of their own activities in the classroom. Digitized training videos can be used to illustrate particular teaching practices. In addition, digitized videos can be used to illustrate norms, standards, and other practices. Digitized video has clear advantages over standard VHS tapes, because it can be linked to systems for access that can include further documentation and support materials. Digitized video can also be used to promote collaborative commentary.

In September 1999, Magdalene Lampert and Jan Hawkins organized an NSF workshop at the University of Michigan that focused on the use of multimodal technology for teacher education. In her workshop presentation, Raven Wallace explained the results of that workshop. The report from the Michigan workshop noted that VHS and digitized video is already being used extensively in teacher training. The report also emphasized the extent to which there has been little attempt to synthesize or document these uses. Currently, there are no standards for these materials, no method of data sharing, and no coordination across efforts. That workshop proposed the "development of a video/ multimedia archive that would allow sharing and recomposition of video and accompanying material." In a sense, the meeting at Carnegie Mellon is a direct response to that suggestion.

Two presentations were specifically dedicated to teacher training materials. Both Richard Lehrer and Lauren Resnick demonstrated completed materials for teacher training that are distributed on CD-ROM. Both sets of materials have video segments joined by a special purpose browser interface. Lehrer's system is designed particularly to illustrate variations in teaching practices from individual teachers; it is currently being used on a limited basis in teacher education. Resnick's system, on the other hand, was designed to inform teachers about optimal practice by illustrating child behaviors that represent high standards for reading and writing achievements of children in kindergarten through fourth grade.

There is no sharp line dividing the use of multimodal materials for teacher training and their use for educational, psychological, linguistic, and sociological research. In fact, all of the participants emphasized the extent to which video documentation of classroom practice can only be done by including teachers as colleagues in the research program. This is an important lesson that suggests that we may well want to focus our collaborative efforts on materials that are of strong interest to teachers, as well as researchers.

Although all of the investigators understood the crucial role of the teacher in generating video observations, many of the other presentations targeted student behaviors more directly. Four of the presentations emphasized the role of cultural diversity in shaping classroom interactions. Sharon Nelson-Barber used video from an Alaskan Inuit classroom to illustrate cooperative learning practices and tolerance for bilingual interactions. Masahiko Minami reported on a series of investigations of narrative styles of parents, children, and teachers in Japanese families in Japan and the United States. Using video, he was able to show how Japanese teachers could teach children narratives in early education programs in the United States. Richard Duran reported on learning in a California Latino community. He found that activities in an after-school computer club program were particularly important in shaping positive attitudes toward learning in children and their parents. Kris Gutierrez used Conversation Analysis methods to study code-switching and peer group value expression in bilingual classrooms in Los Angeles. Together, this group of four studies emphasizes the need for fuller video documentation of crucial peer-peer interactions in groups from diverse cultural backgrounds.

A second area of intense empirical investigations is the study of the collaborative learning of math and science. Presentations by Gordon Wells, Kevin Crowley, Nira Granott, Michelene Chi, Art Graesser, and David Klahr explored these processes using video-based analyses. The focus of Wells' work was the development of a new problem-based learning science curriculum in cooperation with individual teachers. The footage he demonstrated showed a pair of high school students attempting to explain the principles of light diffraction to their classmates using demonstration experiments that were not uniformly successful. Wells interpreted his results in terms of a theory of the spiral of knowing and action theory.   In a very comparable investigation, David Klahr and Eva Toth showed how children tested in controlled conditions in a laboratory setting were able to learn the "control of variables" strategy in all but one variable in experimenting on physical phenomena . They then extended their training materials from the laboratory to the classroom and showed that the resultant learning was richer and equally effective. Kevin Crowley reported on the results of video studies of children visiting science museums with their parents. Crowley found that parents were much more likely to provide mechanistic scientific explanations to their sons, rather than their daughters. Michelene Chi studied children's learning of concepts in physics and biology from various forms of tutorial input. It is now recognized that children learn more from tutors than from conventional instruction, even when the tutors are unskilled. Building on this idea, Art Graesser demonstrated a computer system that replaced the human tutor with an automated computer tutor.

Five other presentations looked at the details of verbal interactions between learners using transcript data. Shoshana Blum-Kulka used transcripts to trace the development of explanations in family and preschool contexts. She documented ways in which teachers continue to shift the rules of the conversation. Clotilde Pontecorvo examined the development of explanation in groups of children asked to describe unfamiliar materials explaining earlier periods in Italian history. Patricia Duff reported on how variations in classroom organization can impact the learning of foreign languages. Tim Koschman emphasized the role of deictic elements and gestures in the six-minute segment from the special issue of *Discourse Processes* mentioned above. Finally, Dick Hayes reported on ways of encouraging students to rewrite faulty prose. He suggested that these same methods could profitably be used to allow students to comment on the effectiveness of teaching methods by annotating video from their teachers.

Data-sharing and Confidentiality

Many of the presentations specifically addressed problems and opportunities in the area of data sharing. Brian MacWhinney and Catherine Snow used the successful example of the Child Language Data Exchange System as an illustration of the ways in which data-sharing can lead scientific advances. Koschman, Wells, Frederiksen, and Graesser echoed this argument. Frederiksen emphasized the need to enrich transcripts with pointers to hierarchically structured models that represent the cognitions underlying specific utterances. He underscored the fact that there are many types of educational discourse and that a focus on classroom discourse would be too restrictive. Like MacWhinney, he stressed the need to share not only transcripts and video, but also coding schemes, analytic frameworks, and other computational tools.

Judith Green pointed out the need for enriching corpora with notes, event maps, essays, diagrams, quizzes, and a wide variety of supporting analyses. The group discussed ways in which such materials can be digitized. Raven Wallace and Richard Lehrer also discussed the importance of linking transcripts to maps and diagrams in their presentations. Green then outlined a social construction view of literacy development in which members of a group are afforded or denied access to opportunities for acquiring a repertoire of literate practices. A group of Australian researchers provided Green and others with multimodal and written data regarding the learning activities of a boy named Simon. Comparing Simon's activities across classes, Green found that Simon was more engaged in the active use of literate practices in his class in Mathematics than in his class in English. Finally, Green turned her attention to problems in the areas of data-sharing, suggesting that recent educational liability cases in California stemming from Proposition 207 have made teachers unwilling to even have visitors enter their classrooms. Similar problems were not reported from other states.

Jane Zuengler devoted much of her presentation to asking fundamental questions about human subjects and confidentiality issues. She noted that many of the protections for Human Subjects currently in force have been stimulated by medical research and are not appropriate for research on teaching. If materials that are used for teacher training are not allowed exemption from human subjects review, then it may become difficult to demonstrate teaching practices. In the end, many of these issues are ones that must be settled by teachers in conjunction with researchers.

Raven Wallace called attention to important data-gathering projects from Fred Goffree and Wil Onk at the Freudenthal Institute in Utrecht, Alan Schoenfeld at Berkeley, and Judith Mumme and Nanette Seago in the Renaissance Project.

The proposed levels of confidentiality protection included in the TalkBank proposal are repeated below as an appendix. There were no specific suggestions for changes to this proposal. However, there was some emphasis on the need for restriction of access to researchers with passwords.

Recommendations

It appears that there are currently five or six major data sets that could be profitably shared within this community and whose owners are willing to make them available. The task of the TalkBank project at this point will be to solidify the rules of data sharing, digitize these data in a uniform and convenient manner, and develop tools that will facilitate collaborative commentary. Rather than attempting to digitize and annotate large corpora, as in the case of CHILDES, it probably makes more sense in the study of classroom discourse to emphasize high-quality digitization and annotation of a much smaller quantity of data that is selected for its interesting and revealing properties.

Once several such data sets are available, the next goal should be the encouragement of collaborative commentary regarding the theoretical and practical issues raised by the interactions. There will be requests for higher quality material with better documentation and more ancillary notes and materials. In addition, we will be able to deal with confidentiality issues within the current small research group, before opening up the discussion (and perhaps the data) to a wider audience.

Current NSF funding for TalkBank is sufficient to support the initial phases of this activity. Therefore, we are not making any recommendations to NSF regarding additional funding at this point. However, we believe that the construction of a shared database of multimodal observations of educational discourse is the best way we have to improve our understanding of the transition of the child to the work force. Therefore, we recommend that NSF continue to support work that is directed to this goal. By the end of 2000, we will again evaluate the status of this work and then provide further recommendations to NSF and other agencies.

NSF Workshop on Classroom Discourse Participants

Steven Bird	University of Pennsylvania	sb@ldc.upenn.edu
Shoshana Blum-Kulka	Tel-Aviv University	mskcusb@mscc.huji.ac.il
Micheline Chi	University of Pittsburgh	chi@vms.cis.pitt.edu
Kevin Crowley	University of Pittsburgh	crowleyk+@pitt.edu
Patricia Duff	Univ. of British Columbia	pduff@unixg.ubc.ca
Ricardo Duran	UC Santa Barbara	duran@education.ucsb.edu
Carl Frederiksen	McGill	ed76@musica.mcgill.ca
Nira Granott	UT Dallas	ngranott@utdallas.edu
Judith Green	UC Santa Barbara	green@education.ucsb.edu
Kris Gutierrez	UC Los Angeles	krisgu@ucla.edu
Arthur Graesser	University of Memphis	a-graesser@memphis.edu
Dick Hayes	Carnegie Mellon	jh50@andrew.cmu.edu
David Klahr	Carnegie Mellon	david.klahr@andrew.cmu.edu
Tim Koschmann	Southern Illinois	tkoschmann@acm.org
Richard Lehrer	University of Wisconsin	rlehrer@facstaff.wisc.edu
Brian MacWhinney	Carnegie Mellon	macw@cmu.edu
Masahiko Minami	San Francisco State	mminami@sfsu.edu
Sharon Nelson-Barber	Wested	snelson@wested.org
Clotilde Pontecorvo	University of Rome	c.pontecorvo@agora.stm.it
Lauren Resnick	University of Pittsburgh	resnick@vms.cis.pitt.edu
Catherine Snow	Harvard University	catherine_snow@harvard.edu
Eva Erdosme Toth	Carnegie Mellon University	etoth@andrew.cmu.edu
Howard Wactlar	Carnegie Mellon University	wactlar@cs.cmu.edu
Johannes Wagner	Odense University	jwa@language.sdu.dk
Gordon Wells	OISE	gwells@oise.utoronto.ca
Raven Wallace	University of Michigan	ravenw@umich.edu
Jane Zuengler	University of Wisconsin	zuengler@facstaff.wisc.edu

Topics Addressed

The participants were given the following list of topics to be addressed:

1. What advice can we give to NSF regarding the general area of children's learning. How can new research improve our understanding of the transition from childhood to the workforce? In what ways will interdisciplinary research help elucidate this issue?

2. How can the study of classroom discourse help us understand the challenges that face particular children from widely divergent backgrounds?

3. Can we compare observations across classrooms to reach generalizations about modes of communication, teaching, and learning. What materials do we need to do this and how can empirical hypotheses be tested using observational data?

4. How can we use observational data to develop standards for learning and teaching?

5. What new computational tools do we need to develop?

6. What new coding schemes do we need to develop and validate?

7. Can multimodal data serve a useful role in teacher training?

8. How can we link studies of classroom interaction to fundamental theoretical issues in sociology, sociolinguistics, and developmental psychology?

9. What are the current audio and video databased resources available to students of classroom discourse and tutorial interactions?

10. What are the prospects for the development of data-sharing between labs or within the community more generally?

11. How can confidentiality and other subject rights be preserved without blocking data-sharing?

12. How can we stimulate collaborative commentary of the type reflected in the recent issue of *Discourse Processes* edited by Tim Koschmann

13. How can we use observational data to reshape thinking in related areas of learning, instruction, and development?

Proposed Levels for Confidentiality Protection

It is proposed that researchers should allow their subjects to determine which of these levels of access are appropriate for both newly collected and older material. In most cases, it will not be necessary to make this full range of options available to subjects, since a particular range, such as one including levels 3, 4, and 5 might be most clearly appropriate.

IRBs should be aware of the complete set of levels and should approve of a researcher's attempt to obtain permission for a certain level.

Level 1:   Data are fully public (public speeches, public interviews, etc.) and generally viewable and copyable over the Internet, although they may still be copyrighted.

Level 2:   Data are open to general viewing and listening by the public across the Internet, but watermarking and other techniques are used to block copying and redistribution.

Level 3:   Transcript data with pseudonyms will be made publicly available. However, the corresponding audio or video data, for which anonymity is more difficult to preserve, will be made available on one of the next six, more restrictive levels.

Level 4:   Data are only available to researchers who have signed a non-disclosure form. This form sets tight standards regarding avoidance of use of personal names when required. It allows some temporary copying or downloading of the data for local analysis, but requires that downloaded files be deleted after a specific period and never further copied or distributed. These requirements are enforced through watermarking and software blocks.

Level 5:   Access is restricted to researchers who have signed non-disclosure forms. In addition, copying is disallowed.

Level 6:   Data viewing requires explicit approval from the contributor of the data. This level would work much like a research laboratory that made copies of videotapes to send to other laboratories and required those laboratories to follow rules about non-distribution of data. However, unlike Level 6, this level would also include mechanisms for insuring that the data would not be copied or distributed.

Level 7:   This level would only allow viewing and listening in controlled conditions under direct on-line supervision. This level is needed for data of a highly personal or revealing nature. This level has been used in the past for the viewing of material from psychiatric interviews.

Level 8:   This level would only allow viewing and listening in controlled conditions under the direct, in person, supervision of the particular researcher. This level is needed for highly sensitive material.

Level 9:   These data would not be viewable, but would be archived in the format of the general system for use by the original investigator only. This level allows the investigator to use the tools of the analysis system without actually "contributing" the data.

This system corresponds closely to procedures currently in use by Human Subjects review committees at the University of Minnesota and the University of California at Berkeley. In addition to protecting subject confidentiality, this system of varying levels can be used to support

the academic interests of the original data collector. For example, if a researcher has not finished publishing the results of a study, access can be set to a more restrictive level. Once the research papers have been published, access can be changed to a less restrictive level.

Some aspects of this system of levels of confidentiality protection can benefit from the development of technical processes. For example, it is possible to create confidentiality by blurring audio and video images. This technology is generally unacceptable for the study of interactional processes, since facial expressions and intonation convey so many important components of communicative meaning. However, there are more sophisticated ways of morphing the face and the voice to images that are still communicatively adequate. Currently, the LDC is using audio morphing to preserve confidentiality in the Corpus of Spoken American English (CSAE) collected by researchers at UC Santa Barbara.   Also, the technique of "watermarking" can be used to prevent or discourage the unauthorized copying of images [89-93].

It would be helpful if IRBs and governmental agencies could clarify their policies regarding the levels of access to older data which did not have this levels designation. Also, it is important to clarify the status of data obtained from participants who are paid as actors and data collected by teachers in their own classrooms with the aim of improving their teaching practices.